2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

Computing common factors of matrix polynomials with applications in system and control theory

Antonio Fazzi[1], Nicola Guglielmi[1] and Ivan Markovsky[2]

***Abstract*— We consider the problem of computing (approximate) greatest common divisors for matrix polynomials and we present some related facts and applications in system and control theory. The main application is to compute the distance of a controllable multi-input multi-output system to the set of uncontrollable ones; then we describe some related results.**
***Index Terms*— Matrix polynomial common factor - Distance to uncontrollability - Behavioral approach**

I. INTRODUCTION

Consider the problem of computing a Greatest Common Divisor (GCD) $L(z)$ for two matrix polynomials $M(z)$ and $N(z)$, which are matrices whose coefficients are polynomials, or equivalently polynomials with matrix coefficients. This is an important topic in the framework of multivariable control systems, and it has been studied by several authors and through different techniques. Some authors find the GCD as a combination of polynomials [1] or transform the block matrix $[M(z)\ N(z)]$ into $[L(z)\ 0]$ [2]. Other methods use the generalized Sylvester matrix [3], [4]. However in practical applications we could deal with coprime matrix polynomials, because the coefficients can be noisy or corrupted, so we need to compute an approximate common factor.

The computation of approximate common factors has been extensively studied in the case of scalar polynomials (e.g. [5], [6], [7], [8], [9], [10], [11], [12], [13], [14] and some references therein) and it is still an active research topic. However, the case of matrix polynomials has not been considered in the scientific literature. The formulation of the problem we consider is to fix the degree d of the sought common factor, and to compute the smallest perturbation on the polynomials coefficients which leads to two matrix polynomials having a common factor of degree d:

$$\begin{aligned} &\text{minimize over} \quad \hat{M}, \hat{N} \quad \|M-\hat{M}\|_F + \|N-\hat{N}\|_F \\ &\text{subject to} \quad \hat{M} = L\bar{M} \quad \hat{N} = L\bar{N} \quad \deg L \geq d \end{aligned} \tag{1}$$

This is a straightforward generalization of the approximate GCD computation for scalar polynomials ([5], [15], [12], [9], [10]). As in the case of scalar polynomials, matrix polynomials having a common divisor define a variety of the Grassman type [16]. In order to have a unique common factor, we should apply some normalization, for example on the leading coefficient.

This work was not supported by any organization.
[1]Antonio Fazzi and Nicola Guglielmi are with Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy antonio.fazzi@gssi.it, nicola.guglielmi@gssi.it
[2]Ivan Markovsky is with the Department ELEC, Vrije Universiteit Brussel (VUB), Pleinlaan 2, Building K, B-1050 Brussels, Belgium imarkovs@vub.ac.be

The approximate greatest common divisor problem is a non-convex optimization problem. To the best of our knowledge, there is no algorithm in the scientific literature for approximate GCD computation for matrix polynomials, so we propose a generalization of the algorithm in [5] (based on a local optimization approach) from scalar to matrix polynomials. We describe in the appendix how the algorithm works, and we use it later on a numerical example.

In the following we present some applications which are related to (approximate) greatest common divisor computation.

II. DISTANCE TO UNCONTROLLABILITY

We present in this section an application of approximate common factors computation for matrix polynomials. It extends the computation of *distance to uncontrollability* from Single Input Single Output systems (presented in [17]) to Multi Input Multi Output systems.

Consider the linear time invariant system $\mathcal{B}(A,B,C,D)$ defined by its state space representation

$$\mathcal{B} = \{(u,y) \mid \dot{x} = Ax + Bu,\ y = Cx + Du\} \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$ are parameters of the system, n denotes the number of states, while m and p are the number of inputs and outputs, respectively. For a given system $\mathcal{B}$ the parameters of its state space representation A, B, C in (2) are not unique: a change of basis $\bar{x} = Vx$ leads to the equivalent representation $\mathcal{B}(VAV^{-1}, VB, CV^{-1}, D)$.

Controllability is a qualitative property of control systems and it is one of the fundamental concepts in modern mathematical control theory. A system described by the representation (2) is said to be controllable if the state function $x(t)$ can be directed from any given state to a desired state in finite time by an input $u(t)$. This classical notion of controllability [18] for the system (2) is related to the possibility of changing the state of the system, hence it is only a property of the parameters A and B; from the mathematical point of view it involves a rank test on the so called controllability matrix

$$\mathcal{C}(A,B) = \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}. \tag{3}$$

In particular it holds true that the system (2) is controllable if and only if the matrix (3) is full rank.

Asking whether a system is controllable or not is a yes-no answer involving the computation of the numerical rank of matrix (3). However arbitrary small perturbation on the

parameters of the system can switch its controllability property. Nearly uncontrollable systems are associated with ill conditioned computational problems (e.g. [19], [20]). Moreover a nearly uncontrollable system can signal fundamental trouble with the mathematical model or even the underlying physical system [21]. Because of this fact, it is useful to have a measure which defines how far a controllable system is from the nearest uncontrollable system. This leads to the problem of computing the *distance to uncontrollability*.

The definition of distance to uncontrollability considered in the scientific literature [22] is only a property of the parameters A and B of the state representation of the system. It is defined as the structured distance to singularity of the matrix $\mathcal{C}(A,B)$ in (3), so it involves the computation of the smallest pair (w.r.t. the Frobenius norm) $(\Delta A, \Delta B)$ so that

$$\mathcal{C}(A,B) + \mathcal{C}(\Delta A, \Delta B)$$

is singular. It was shown by Eising [23], [24] that this is equivalent to the following global optimization problem in two real variables:

$$\min_{\lambda \in \mathbb{C}} \sigma_n (A - \lambda I \quad B) \tag{4}$$

where σ_n denotes the n-th (the smallest) singular value. Since these definitions depend on the state space representation, they are not invariant because of the nonuniqueness of the parameters A, B (anyway for physical state space descriptions it is the fixed parameters which are important). In order to avoid these issues we adopt the behavioral setting [25], [26], [27], where the notion of controllability and the associated distance problems become a (representation invariant) property of the system and not of the parameters we choose for its representation. In this framework, the dynamical system (2) is viewed as the set of its trajectories. The controllability property is described by the possibility of concatenating any two trajectories, up to a delay of time.

Definition 1: Let $\mathcal{B}$ be a linear time invariant dynamical system, which is a set of trajectories (vector valued functions of time). $\mathcal{B}$ is said to be controllable if for all $w_1, w_2 \in \mathcal{B}$ there exists a $T > 0$ and a $w \in \mathcal{B}$ such that

$$w(t) = \begin{cases} w_1(t) & \text{for } t < 0 \\ w_2(t) & \text{for } t \geq T \end{cases}$$

A system is uncontrollable if it is not controllable.

The controllability property (in the behavioral setting) involves a rank test on a particular representation of the system. Any linear time invariant system can be described by a kernel representation [28]; hence given the system $\mathcal{B}$, there exists a matrix polynomial $R(z) \in \mathbb{R}^{p \times (m+p)}[z]$ such that

$$\mathcal{B}(R) = \{ w \mid R_0 w + R_1 \sigma w + \cdots + R_l \sigma^l w = 0 \}, \tag{5}$$

where σ is the shift operator (in the discrete case) or the derivative operator (in the continuous case). The controllability property is related to the rank of the matrix polynomial $R(z)$, and in particular we have the following Lemma [25]:

Lemma 2: The system $\mathcal{B}$ is controllable (according to Definition 1) if and only if the matrix polynomial

$$R(z) = R_0 + R_1 z + \cdots + R_l z^l$$

is left prime, i.e., $R(z)$ is full row rank for all $z \in \mathbb{C}$.

Alternatively to (5), a Multi Input Multi Output linear time invariant system can be represented by its input/output representation

$$\mathcal{B}_{i/o}(P,Q) = \left\{ \begin{pmatrix} u \\ y \end{pmatrix} \,\middle|\, P(z)y = Q(z)u \right\}.$$

This is done, starting from the kernel representation (5), by splitting the variables w into two sets (the inputs u and the outputs y) and partitioning the columns of $R(z) = (Q(z) \; P(z))$ accordingly.

As a consequence of Lemma 2 we have the following result [29]:

Corollary 3: The system $\mathcal{B}$ is controllable if and only if the matrix polynomials $P(z)$ and $Q(z)$ are (left) coprime. The presence of left common factors in $P(z)$ and $Q(z)$ leads to loss of controllability.

Proof: Because of Lemma 2 we show that $P(z)$ and $Q(z)$ have no (left) common factors if and only if the block matrix $(P(z) \; Q(z))$ is full row rank $\forall z$. Assume that P and Q have no left common factors, so they are left coprime. Hence [2] there exists a unimodular matrix U [1] such that

$$(P(z) \; Q(z))U(z) = (I \; 0),$$

so the matrix is full rank for all z.

On the other hand, suppose $P(z) = C(z)\bar{P}(z), Q(z) = C(z)\bar{Q}(z)$. We can factorize the block matrix $(P(z) \; Q(z))$ as $C(z)(\bar{P}(z) \; \bar{Q}(z))$. The zeros of $C(z)$ are the only points which makes the matrix $(P(z) \; Q(z))$ lose rank, since $\bar{P}, \bar{Q}$ are coprime.

Hence the system $\mathcal{B}$ is controllable if and only if the matrices $P(z)$ and $Q(z)$ have no left common factors of degree one or more. ∎

Let $\mathcal{L}_{uc}$ be the set of uncontrollable linear time-invariant systems with $m \geq 1$ inputs and $p \geq 1$ outputs,

$$\mathcal{L}_{uc} = \{ \mathcal{B} \mid \mathcal{B} \text{ uncontrollable LTI system} \}$$

and define the distance between two arbitrary systems by

$$\text{dist}(\mathcal{B}(P,Q), \mathcal{B}(\bar{P},\bar{Q})) = \| (P \quad Q) - (\bar{P} \quad \bar{Q}) \|_F, \tag{6}$$

where, with an abuse of notation, we identified the matrix polynomials by the vectors containing their coefficients[2]. The problem of computing the distance to uncontrollability is the following:

Problem 4: Given a controllable system $\mathcal{B}(P,Q)$, find

$$d(\mathcal{B}) = \min_{\bar{\mathcal{B}} \in \mathcal{L}_{uc}} \text{dist}(\mathcal{B}, \bar{\mathcal{B}}),$$

[1] A unimodular matrix is a matrix polynomial whose determinant is a non zero constant.

[2] The parameters P and Q which identify the system are not unique, but an equivalent representation can be obtained multiplying both P and Q by the same unimodular matrices. In order to have a well posed definition of distance we need to apply some normalization, e.g. we can assume P to be monic; this assumption is not general though.

where the distance is the one defined in (6).

Problem 4 is an approximate left common factor computation problem (1) with $d = 1$.

In the following, we propose a numerical example in order to show the benefits of our approach (the behavioral setting) and the performances of our algorithm (the one we briefly illustate in the Appendix; it generalizes the local optimization method proposed in [5], based on integration of a system of ODEs describing the gradient system associated with a suitable functional). Problem 4 could be stated in the same way it was developed for Single-Input Single-Output systems; the novelty consists in the algorithm we have for its numerical solution (there is no algorithm in the scientific literature for computing approximate common factors of matrix polynomials, to the best of our knowledge).

A. Numerical examples

Consider the following system given by its state space representation:

$$\begin{aligned}&\mathcal{B}(A,B,C,D)\\ &A=\begin{pmatrix}1 & 1 & 1 & 0\\ 0.1 & 3 & 5 & 0\\ 0 & -1 & -1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}, B=\begin{pmatrix}1 & 2\\ 0.1 & 0.5\\ 0 & 1\\ \ell & \ell\end{pmatrix},\\ &C=\begin{pmatrix}1 & 2 & 0.1 & 0.3\\ 3 & 0.1 & 0.1 & 0.5\end{pmatrix}, D=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}.\end{aligned} \tag{7}$$

Observe the particular structure of the system. If we set $\ell = 0$, the system is uncontrollable independently of all the other values (the last row of the controllability matrix (3) is $\ell[1, 1, \ldots, 1]$). Arbitrary perturbations on the parameter ℓ make the system controllable (again check the rank of the controllability matrix (3) for $\ell \neq 0$), so modifying the value of ℓ we are able to build a controllable system which is arbitrarily close to an uncontrollable system.

If we want to move to the behavioral setting, we need a way to switch from the state representation (7) to the kernel representation

$$\mathcal{B} = \{w | R_0 w + R_1 \sigma w + \cdots + R_r \sigma^r w = 0\}.$$

In order to compute the matrix polynomial $R(z)$, we can write the system (2) in extended form (discretizing and rewriting the equations):

$$\begin{aligned}\begin{pmatrix}y(k)\\ y(k+1)\\ y(k+2)\end{pmatrix} &= \begin{pmatrix}C\\ CA\\ CA^2\end{pmatrix} x(k)\\ &+ \begin{pmatrix}D & & \\ CB & D & \\ CAB & CB & D\end{pmatrix}\begin{pmatrix}u(k)\\ u(k+1)\\ u(k+2)\end{pmatrix}.\end{aligned} \tag{8}$$

Since we want to eliminate the state x in order to get the matrices P and Q of the input/output representetion, we can premultiply the equation (8) by the matrix in the left null space of the observability matrix $(C; CA; CA^2)$:

$$\begin{aligned}(P_2, P_1, P_0)\begin{pmatrix}y(k)\\ y(k+1)\\ y(k+2)\end{pmatrix} &= (P_2, P_1, P_0)\begin{pmatrix}C\\ CA\\ CA^2\end{pmatrix} x(k)\\ +(P_2, P_1, P_0)\begin{pmatrix}D & & \\ CB & D & \\ CAB & CB & D\end{pmatrix}&\begin{pmatrix}u(k)\\ u(k+1)\\ u(k+2)\end{pmatrix}\end{aligned} \tag{9}$$

We denote by $P(z) = P_0 + P_1 z + P_2 z^2$, while $Q(z)$ is given by the coefficients in the matrix product on the last row of (9):

$$\begin{aligned}Q_2 &= P_2 D + P_1 CB + P_0 CAB;\\ Q_1 &= P_1 D + P_0 CB;\\ Q_0 &= P_0 D.\end{aligned}$$

In this way we have a kernel representation for the system in (7).

If we set $\ell = 0$, the system (7) is uncontrollable, and the two polynomials which give the kernel representation are the following (all the numerical values are rounded to three decimal places):

$$\begin{aligned}P(z) &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} z^2 + \begin{pmatrix}0.516 & -0.064\\ 0.173 & -0.460\end{pmatrix} z\\ &+ \begin{pmatrix}-0.700 & 0.670\\ 0.242 & -0.232\end{pmatrix}\\ Q(z) &= \begin{pmatrix}1.434 & -1.092\\ 0.328 & -0.135\end{pmatrix} z^2 + \begin{pmatrix}-1.378 & 1.142\\ -0.912 & 0.487\end{pmatrix} z\\ &+ \begin{pmatrix}0.249 & -0.267\\ 0.739 & -0.463\end{pmatrix}.\end{aligned} \tag{10}$$

We can check that the two matrices $P(z)$ and $Q(z)$ in (10) have as common factor $\begin{pmatrix}0 & 1\\ z-1 & 0.509\end{pmatrix}$, and the associated Sylvester resultant $S(P, Q)$ (according to the classical definition presented in the literature [30], [31]) is rank deficient with co-rank 1 (observe the correspondence among the co-rank of the controllability matrix (3), the co-rank of the Sylvester resultant and the degree of the determinant of the common factor). If we perturb the value of ℓ in (7), the correspondng system is controllable (the controllability matrix (3) is full rank) and the controllability is reflected also in the behavioral sense, since the corresponding matrix polynomials $\hat{P}, \hat{Q}$ coming from the kernel representation are coprime (the associated Sylvester resultant $\mathcal{S}(\hat{P}, \hat{Q})$ is full rank). However if we consider the problem of computing the distance from a controllable system to the closest uncontrollable system, the values of the distance problems are different.

In the following experiment (Table I) we consider the system in (7) and we list some values of the computed *distance to uncontrollability* corresponding to different perturbations of the parameter ℓ. The distance in the classic sense is computed by looking for the numerical solution of (4). In [32] are listed several problems arising by solving the non convex global optimization problem (4) by local optimization approaches. The algorithm we use for solving

(4) is the one proposed in [33], where the authors underline there are no standard methods for computing distance to uncontrollability, which is a more difficult problem than other distance problems, such as distance to singularity or distance to instability.

The distance in the behavioral sense involves the computation of an approximate left common factor between $\hat{P}(z)$ and $\hat{Q}(z)$, the two matrix polynomials coming from the input/output representation, which are coprime since the corresponding system is controllable (the algorithm we use for computing these values of distance is the one briefly summarized in the Appendix).

We consider then the converse problem: we start from an uncontrollable system in the behavioral sense (i.e., two matrix polynomials having a left common factor) and we switch, after a perturbation on the polynomials, to the state space representation in order to analyze the values of the distance to the closest uncontrollable system.

Consider again the polynomials in (10). We switch from the kernel to the state space representation calling the function *lmf2ss* from the Polyx Toolbox [34]. The computed representation is:

$$\mathcal{B} = \mathcal{B}(A, B, C, D)$$
$$A = \begin{pmatrix} 1.018 & 1 & -0.438 & 0 \\ -0.191 & 0 & 0.567 & 0 \\ 0.646 & 0 & 1.932 & 1 \\ -0.542 & 0 & -1.010 & 0 \end{pmatrix} D = \begin{pmatrix} 0 & 2 \\ -3.040 & 7.90 \end{pmatrix}$$
$$B = \begin{pmatrix} 0.330 & -0.374 \\ -1.364 & 0.808 \\ -3.769 & 2.676 \\ -1.714 & -0.769 \end{pmatrix} C = \begin{pmatrix} 0 & 0 & 2 & 0 \\ -3.040 & 0 & 7.90 & 0 \end{pmatrix}. \tag{11}$$

Since the system is uncontrollable by construction, the associated controllability matrix (3) is singular, as expected.

We now add some random perturbations to the coefficients of the polynomials, in order to see how the corresponding state space representation changes (in terms of *distance to uncontrollability*). The results are given in Table II.

This last experiment confirm the results we observed in Table I, that is the distances in the bahvioral setting are smaller than the ones in the classic sense, so the behavioral approach seems to be more accurate in finding which is the closest uncontrollable system. We notice the difference in the computed values of the distance in the classic sense comparing Table I and Table II, as a consequence of the different parameters we adopted for the state space representation. We do not observe such a variation in the distances in the behavioral sense, where the computed values reflect the perturbations on the data.

All the previous numerical experiments are online available [35] and they can be reproduced. In the first example the perturbation is deterministic, so the numerical results are exactly the same; on the other side we remark that the perturbation in the second example is randomly generated, so we cannot expect to get exactly the same results but only similar ones.

TABLE I

DISTANCE TO UNCONTROLLABILITY FOR A PERTURBED STATE SPACE REPRESENTATION AND THE CORRESPONDING KERNEL REPRESENTATION

ℓ	d. uncontr. (classic)	d. uncontr. (behavioral)
10^{-4}	$5.4 \cdot 10^{-5}$	$9 \cdot 10^{-9}$
10^{-3}	$5.4 \cdot 10^{-4}$	$6.6 \cdot 10^{-6}$
10^{-2}	$5.4 \cdot 10^{-3}$	$6.6 \cdot 10^{-5}$

TABLE II

DISTANCES TO UNCONTROLLABILITY OF A PERTURBED KERNEL REPRESENTATION AND THE CORRESPONDING STATE SPACE REPRESENTATION

norm perturbation	d. uncontr. (behavioral)	d. uncontr. (classic)
$3.98 \cdot 10^{-5}$	$3 \cdot 10^{-6}$	0.120
$4.38 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$	0.119
$4.7 \cdot 10^{-3}$	$4.5 \cdot 10^{-4}$	0.119

III. OTHER APPLICATIONS INVOLVING COMMON FACTORS COMPUTATION

A. Controllable and uncontrollable subsystems

In section II we described how to compute the distance of a controllable linear time invariant system from the uncontrollable ones. Once we compute a solution of Problem 4, we can split the system into two parts, controllable and uncontrollable [28]. A nice way to define the controllable part of a system is via subsystems. If $\mathcal{B}_1, \mathcal{B}_2$ are two dynamical systems, we say $\mathcal{B}_1$ is a subsystem of $\mathcal{B}_2$ if the set of trajectories of the first system is a subset of the ones of the second system. This can be expressed also in terms of kernel representation: if $R_1(z), R_2(z)$ are the kernel representations of $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively, then $\mathcal{B}_1$ is a subsystem of $\mathcal{B}_2$ if there exist a matrix polynomial $F(z)$ such that $R_2(z) = F(z)R_1(z)$.

Definition 5: The controllable part of a system $\mathcal{B}$, denoted by $\mathcal{B}_c$, is defined as the largest controllable linear time-invariant subsystem of $\mathcal{B}$.

Assume that the system $\mathcal{B}$ is defined by the difference equation $R(z)w = 0$ for a certain matrix polynomial $R(z) \in \mathbb{R}^{g \times q}$, and we are interested in computing $\mathcal{B}_c$. If the matrix R can be factored as $R(z) = C(z)\hat{R}(z)$, with $C(z) \in \mathbb{R}^{g \times g}$, $\det(C(z)) \neq 0$, $\hat{R}(z) \in \mathbb{R}^{g \times q}$, then the system represented by the difference equation $\hat{R}(z)w = 0$ is the controllable part of $\mathcal{B}$ (assuming $\hat{R}(z)$ is left prime). If such a factorization does not exist, we could consider an approximate common factor computation between the two blocks $P(z), Q(z)$ of the matrix $R(z)$ in the input/output representation.

B. Minimal kernel representation

A further problem involving GCD computation is to compute a minimal kernel representation starting from a non-minimal one [25]. A non-minimal kernel representation is a vector $\tilde{R}(z)$ whose elements are $g_i \times q$ matrix polynomials, such that the trajectories w of the dynamical system

satisfy the difference equation $\tilde{R}(\sigma)w = 0$. The kernel representation is minimal if $\tilde{R}(z)$ has no redundant rows. In the matrix autonomous case we can compute a minimal kernel representation by looking at the GCD of the elements $R_i(z) \in \mathbb{R}^{g_i \times q}[z]$ of $\tilde{R}$, i.e.,

$$R_{\min}(z) = \text{GCD } (R_1(z), \ldots, R_l(z)). \tag{12}$$

C. Intersection of behaviors

The *common dynamics estimation* problem has been presented in [17] in the case of scalar polynomials. We reformulate it in the framework of matrix polynomials.

Problem 6: Given a set of N multivariable autonomous linear time invariant systems $\mathcal{B}_1, \ldots, \mathcal{B}_N$ (with the same number of variables), find their common dynamic, defined as the intersection of the systems $\mathcal{B}_1, \ldots, \mathcal{B}_N$, i.e.,

$$\mathcal{B} = \mathcal{B}_1 \cap \cdots \cap \mathcal{B}_N.$$

Consider for each system its kernel representation

$$\mathcal{B}_i = \ker(P_i(\sigma)) := \{w \mid P_0 w + P_1 \sigma w + \cdots + P_n \sigma^n w = 0\}. \tag{13}$$

where each P_i is a $g_i \times q$ matrix polynomial. Problem 6 is equivalent to a greatest common divisor computation for the matrices $P_1, \ldots, P_N$. In the case there is no an exact common factor we can consider the problem of computing an approximate common factor. The following illustration can be extended from 2 to any number N of systems.

$$\mathcal{B}_1 = \ker(R_1), \quad \mathcal{B}_2 = \ker(R_2),$$

$$\mathcal{B} = \mathcal{B}_1 \cap \mathcal{B}_2 = \ker \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} = \ker \Big(GCD(R_1, R_2) \Big).$$

IV. CONCLUSION

Given a controllable linear time invariant Multi-Input Multi-Output system, we proposed a representation invariant measure for computing the *distance to uncontrollability*. In the proposed framework (the behavioral approach) the controllability property is regarded as a property of the whole system, and not only of the (non unique) parameters choosen for its state space representation. The problem is restated as an approximate left common factor computation between two matrix polynomials; we have now an algorithm for the numerical solution of this problem (whose ideas are briefly summarized in the Appendix). More applications related to GCDs computation for matrix polynomials are: decomposition of a given system into controllable and uncontrollable subsystems, converting a kernel representation into a minimal one and finding the intersection of linear time-invariant systems.

APPENDIX: AN ALGORITHM FOR COMPUTING APPROXIMATE COMMON FACTORS OF MATRIX POLYNOMIALS

We describe here the ideas behind the algorithm for computing the numerical solution of Problem 4, i.e., the computation of approximate left common factors for two matrix polynomials. It generalizes the algorithm presented in [5] from scalar to matrix polynomials.

The starting data are two left coprime matrix polynomials $P(z), Q(z)$ (having the same number of rows) and the associated modified Sylvester resultant $S_w(P, Q)$ (as defined in [30]), which is full rank. A rough starting estimate for the distance to uncontrollability can be the smallest singular value of $S_w(P, Q)$; however the sought solution is given by the structured distance to singularity. The goal is to perturb the coefficients of the matrix polynomials (in a minimal way) till the associated Sylvester resultant becomes singular. This is done by iteratively adding a structured perturbation of the form ϵE, where $\epsilon \in \mathbb{R}$ is the norm of such a perturbation, while E is a Sylvester matrix of (Frobenius) norm 1 which identifies as ϵE the minimizer of the smallest singular value of $S_w + \epsilon E$ over the ball of matrices whose norm is at most ϵ. In this way we can move E and ϵ separately on two different levels:

- at the inner level we fix the value of ϵ and we aim at minimizing the smallest singular value σ of $S_w + \epsilon E$ by looking for the stationary points of the associated gradient system;
- at the outer level we know the value of E and we want to move the value of ϵ in order to find the smallest perturbation which vanishes the value of σ.

During the inner iteration we look for minimizing the value of σ for a fixed value of ϵ. Hence we need to compute an optimal perturbation E which minimizes the smallest singular value σ of $S_w + \epsilon E$ over the set of matrices of unit Frobenius norm. This is done by looking for a smooth path of matrices $E(t)$ (of unit Frobenius norm) along which the value of σ is decreasing. This direction is computed through the stationary points of the following gradient system (omitting the time dependence)

$$\dot{E} = -P_{\mathcal{S}}(uv^\top) + \langle P_{\mathcal{S}}(uv^\top), E \rangle E \tag{14}$$

where $P_{\mathcal{S}}(\cdot)$ is the operator which project the argument onto the Sylvester structure, while u, v are the left and right singular vectors associated to σ. The system of ODEs (14) comes from the combination of the expression of the derivative of eigenvalues, the relation between eigenvalues and singular values of a matrix and the solution of a constrained minimization problem (the reader can check [36] for further details and the formula for the projection $P_{\mathcal{S}}$). It can be proved that the singular value σ associated to u and v is actually decreasing along the solution of (14), and it reaches the point of minimum in correspondence to the stationary points of the equation (14).

Once we have computed the direction E and the corresponding singular value, denoted by $\sigma(\epsilon)$, we need to move the value of ϵ in order to find the smallest perturbation to the original Sylvester matrix which decreases, up to a small tolerance, the smallest singular value $\sigma(\epsilon)$ of the matrix $S_w + \epsilon E$. This can be done by a root finding algorithm (e.g. the Newton method), possibly coupled with a bisection step, in order to look for possible better solutions once $\sigma(\epsilon)$ reaches the fixed tolerance.

A more detailed description of the algorithm, its main properties and features and the similarities and differences

with respect to the scalar case, can be found in [36]. We remark that in [36] it is assumed (for the sake of simplicity) that all the common factors have a full rank leading coefficient, while in this paper we used common factors with rank deficient leading coefficients; this is done by changing the functional to be minimized in a suitable way.

ACKNOWLEDGMENT

N. G. thanks the Italian INdAM GNCS (Gruppo Nazionale di Calcolo Scientifico) for financial support. I. M. received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013) / ERC Grant agreement number 258581 "Structured low-rank approximation: Theory, algorithms, and applications" and Fund for Scientific Research Vlaanderen (FWO) projects G028015N "Decoupling multivariate polynomials in nonlinear system identification" and G090117N "Block-oriented nonlinear identification using Volterra series"; and Fonds de la Recherche Scientifique (FNRS) – FWO Vlaanderen under Excellence of Science (EOS) Project no 30468160 "Structured low-rank matrix / tensor approximation: numerical optimization-based algorithms and applications". All the authors thank the reviewers for their useful comments and suggestions.

REFERENCES

[1] C. C. Mac Duffee, *The Theory of Matrices*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1933.

[2] W. A. Wolovich, *Linear Multivariable Systems*, ser. Applied Mathematical Sciences. New York, NY: Springer New York, 1974, vol. 11.

[3] B. Anderson and E. Jury, "Generalized Bezoutian and Sylvester matrices in Multivarioable Linear Control," *IEEE Trans. Automat. Contr.*, vol. 21, pp. 551–556, 1976.

[4] R. R. Bitmead, S. Y. Kung, B. D. Anderson, and T. Kailath, "Greatest Common Divisors via Generalized Sylvester and Bezout Matrices," *IEEE Trans. Automat. Contr.*, vol. 23, pp. 1043–1047, dec 1978.

[5] A. Fazzi, N. Guglielmi, and I. Markovsky, "An ODE-based method for computing the approximate greatest common divisor of polynomials," *Numer. Algorithms*, vol. 81, pp. 719–740, 2019.

[6] D. A. Bini and P. Boito, "Structured matrix-based methods for polynomial ϵ-gcd: Analysis and comparisons," in *Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation*, ser. ISSAC '07. New York, NY, USA: ACM, 2007, pp. 9–16.

[7] P. Chin, R. M. Corless, and G. F. Corliss, "Optimization strategies for the approximate gcd problem," in *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation*, ser. ISSAC '98. New York, NY, USA: ACM, 1998, pp. 228–235.

[8] N. Karcanias, S. Fatouros, M. Mitrouli, and G. Halikias, "Approximate greatest common divisor of many polynomials, generalised resultants, and strength of approximation," *Computers & Mathematics with Applications*, vol. 51, no. 12, pp. 1817–1830, 2006.

[9] I. Markovsky and S. Van Huffel, "An algorithm for approximate common divisor computation," in *Proc. of the 17th Symp. on Math. Theory of Networks and Systems*, 2006, pp. 274–279.

[10] W. Qiu, Y. Hua, and K. Abed-Meralm, "A subspace method for the computation of the GCD of polynomials," *Automatica*, vol. 33, pp. 741–743, apr 1997.

[11] A. Terui, "GPGCD: an iterative method for calculating approximate GCD of univariate polynomials," *Theor. Comput. Sci.*, vol. 479, pp. 127–149, 2013.

[12] K. Usevich and I. Markovsky, "Variable projection methods for approximate (greatest) common divisor computations," *Theor. Comput. Sci.*, vol. 681, pp. 176–198, 2017.

[13] C. J. Zarowski, X. Ma, and F. W. Fairman, "Qr-factorization method for computing the greatest common divisor of polynomials with inexact coefficients," *IEEE Trans. Signal Processing*, vol. 48, pp. 3042–3051, 2000.

[14] Z. Zeng and B. H. Dayton, "The approximate gcd of inexact polynomials," in *Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation*, ser. ISSAC '04. New York, NY, USA: ACM, 2004, pp. 320–327.

[15] L. Zhi and Z. Yang, "Computing approximate GCD of univariate polynomials by structure total least norm," in *D. Wang, L. Zhi (eds.) Int. Workshop on Symbolic-Numeric*, Xian, China, 2005, pp. 188–201.

[16] N. Karcanias and G. Halikias, "Approximate zero polynomials of polynomial matrices and linear systems," *Linear Algebra Appl.*, vol. 439, pp. 1091–1103, 2013.

[17] I. Markovsky, A. Fazzi, and N. Guglielmi, "Applications of Polynomial Common Factor Computation in Signal Processing," in *Latent Variable Analysis and Signal Separation. Lecture Notes in Computer Science*. Springer, 2018, pp. 99–106.

[18] R. Kalman, "On the general theory of control systems," *IFAC Proceedings Volumes*, vol. 1, no. 1, pp. 491 – 502, 1960, 1st International IFAC Congress on Automatic and Remote Control, Moscow, USSR, 1960.

[19] A. Laub, "Numerical linear algebra aspects of control design computations," *IEEE Trans. Automat. Contr.*, vol. 30, no. 2, pp. 97–108, February 1985.

[20] J. W. Demmel, "On condition numbers and the distance to the nearest ill-posed problem," *Numer. Math.*, vol. 51, pp. 251–290, 1987.

[21] R. Byers, "Detecting nearly uncontrollable pairs," in *Numerical methods proceedings of the international symposium MTNS-89, volume III*. Springer-Verlag, 1990, pp. 447–457.

[22] C. Paige, "Properties of numerical algorithms related to computing controllability," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 130–138, February 1981.

[23] R. Eising, *The distance between a system and the set of uncontrollable systems*, ser. Memo COSOR. Technische Hogeschool Eindhoven, 1982.

[24] ——, "Between controllable and uncontrollable," *Systems Control Lett.*, vol. 4, pp. 263–264, 1984.

[25] J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory*, ser. Texts in Applied Mathematics. New York, NY: Springer New York, 1998, vol. 26.

[26] J. C. Willems, "The behavioral approach to open and interconnected systems: Modeling by tearing, zooming, and linking," *IEEE Control Syst. Mag.*, vol. 27, pp. 46–99, 2007.

[27] I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor, *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. SIAM, 2006.

[28] J. C. Willems, "From time series to linear system. Part I. Finite dimensional linear time invariant systems," *Automatica*, vol. 22, pp. 561–580, sep 1986.

[29] ——, "A framework for the study of dynamical systems." in *Kaashoek M.A., van Schuppen J.H., Ran A.C.M. (eds) Realization and Modelling in System Theory. Progress in Systems and Control Theory, vol 3.* Birkhäuser Boston, 1990, pp. 43–59.

[30] I. Gohberg and G. Heinig, "The Resultant Matrix and its Generalizations. I. The Resultant Operator for Matrix Polynomials," in *Convolution Equations and Singular Integral Operators*. Basel: Birkhäuser Basel, 2010, pp. 65–88.

[31] I. Gohberg, M. A. Kaashoek, and L. Lerer, "The resultant for regular matrix polynomials and quasi commutativity," *Indiana U. MATH. J.*, vol. 57, pp. 2793–2814, 2008.

[32] M. Gu, "New mwthods for estimating the distance to uncontrollability," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 989–1003, 2000.

[33] J. V. Burke, A. S. Lewis, and M. L. Overton, "Pseudospectral components and the distance to uncontrollability," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 350–361, 2004.

[34] *Polyx: The Polynomial Toolbox for MATLAB, Version 1.5*, Polyx, Ltd., Praha, Czech Republic, 1997.

[35] A. Fazzi, N. Guglielmi, and I. Markovsky, "Experiments in the paper "computing common factors of matrix polynomials with applications in system and control theory," https://github.com/fazziant/Computational-methods/tree/Approximate-common-factors-matrix-polynomials, 2019.

[36] ——, "Computing approximate common factors of matrix polynomials," submitted, https://arxiv.org/abs/1907.13101.